MISB MOTION IMAGERY STANDARDS BOARD

STANDARD

Time Stamping and Transport of Compressed Motion Imagery and Metadata

MISB Standard 0604.2

9 June 2011

1 Scope

This Standard defines methods to time stamp compressed motion imagery streams and to transport compressed motion imagery and metadata in an MPEG-2 Transport Stream protocol.

Within the DoD, the MPEG-2 Transport Stream is generically called “Xon2” where “X” defines existing or future video compression technologies and “on2” refers to the use of MPEG-2 transport streams and files. Xon2 provides a migration path to inject improved compressions technologies, which will yield improved image quality and/or reduced bandwidths. Currently, MPEG-2 video compression and H.264/AVC motion imagery compression are fielded technologies that leverage the Xon2 construct for carriage of motion imagery, audio, and metadata. Time stamping of MPEG-2 and H.264/AVC are covered in this standard. This Standard does not address audio and metadata time stamping.

2 References

Normative References

[1] ISO/IEC 13818-1:2007, Information technology - *Generic coding of moving pictures and associated audio information: Systems*

[2] Motion Imagery Standards Profile 6.2, *DoD/IC/NSG Motion Imagery Standards Board, Jun 2011* http://www.gwg.nga.mil/misb/misppubs.html

[3] MISB RP 0101.1, *Use of MPEG-2 System Streams in Digital Motion Imagery Systems, Jan 2011*

[4] MISB STANDARD 0603.1, *Common Time Reference for Digital Motion Imagery using Coordinated Universal Time (UTC), Jun 2011*

[5] MISB STANDARD 0605.3, *Inserting Time Code and Metadata in High Definition Uncompressed Video, Jun 2011*

[6] ISO/IEC 13818-2:2000, Information technology - *Generic coding of moving pictures and associated audio information: Video*

[7] MISB STANDARD 9715, *Time Reference Synchronization*

[8] SMPTE ST 12-1:2008, Television - *Time and Control Code*

[9] SMPTE ST 328:2000, Television - *MPEG-2 Video Elementary Stream Editing Information*

[10] SMPTE ST 309:1999, Television - *Transmission of Date and Time Zone Information in Binary Groups of Time and Control Code*

[11] ISO/IEC 14496-10:2009, Information technology - *Coding of audio-visual objects - Part 10: Advanced Video Coding*

[12] SMPTE ST 336:2007, *Data Encoding Protocol using Key-Length-Value*

[13] SMPTE RP 217:2001, *Nonsynchronized Mapping of KLV Packets into MPEG-2 System Streams*

Informative References

[14] SMPTE EG 40:2002, *Conversion of Time Values Between SMPTE 12M Time Code, MPEG-2 PCR Time Base and Absolute Time*

3 Definitions

Synchronous Metadata Multiplex Method: metadata multiplexing into a MPEG-2 Transport Stream in accordance with ISO/IEC 13818-1 [1]

Asynchronous Metadata Multiplex Method: metadata multiplexing into a MPEG-2 Transport Stream in accordance with SMPTE RP-217 [13]

4 Introduction

The MPEG-2 transport layer (ISO 13818-1 [1]) provides an infrastructure for the carriage of motion imagery, audio and metadata in a single motion imagery stream as shown in Figure 4-1.

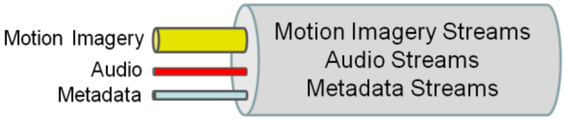

Figure 4-1: MPEG-2 Transport Stream (media essences)

The Motion Imagery Standards Profile (MISP) [2] endorses the use of MPEG-2 Transport Stream [1, 3] (denoted “Xon2” for compression standards consistent with the MPEG-2 Transport Stream). For example, use of MPEG-2 Transport Stream as a means for carriage of H.264 compressed motion imagery is defined in [1].

Motion imagery analysis and processing applications depend on motion imagery and metadata that are temporally correlated. Time stamps, when done accurately and applied to both the motion imagery and the metadata at the point of collection provide a means to ensure that the motion imagery and the metadata can maintain correlation. MISB STD 0603[4] defines the use of UTC as the deterministic common time reference for correlation of motion imagery frames and metadata. STD 0603 also defines two time stamps used in motion imagery: a Commercial Time Stamp and a Precision Time Stamp. These time stamps are referenced herein.

This standard provides guidance to accurately time stamp MPEG-2 and H.264 compressed motion imagery using Coordinated Universal Time (UTC). The advantages of using UTC as the master clock reference for video and metadata are outlined in [4], which also describes various time formats and the relationships between them.

This standard identifies how to insert both a Commercial Time Stamp and a Precision Time Stamp into MPEG-2 and H.264 compressed motion imagery. The Commercial Time Stamp, used widely in the broadcast community, is a relative time stamp that facilitates editing, asset management, and post processing functions. Because of its relative nature, and that it might be modified within the processing workflow, Commercial Time Stamp is not recommended as a persistent time stamp for correlating motion imagery with metadata.

This standard provides guidance for the multiplexing of motion imagery and metadata within the MPEG-2 Transport Stream protocol [1] with consideration of synchronized component essence alignment at the transport level for presentation/display.

This standard does not address metadata time stamping specifically, but assumes that time stamps are properly inserted into metadata where specified per existing MISB metadata set specific documents.

NOTE: For time stamps to be completely meaningful they must be specified with known accuracy in relation to both the imagery point of capture and the metadata point of capture. The qualification of time stamps accuracy is not well characterized presently in the MISB standards, but it is a future objective of the MISB to provide such qualifications.

5 Time Stamping Compressed Motion Imagery

5.1 Timing Information Source

Time stamps are introduced into a compressed motion imagery stream in one of the following ways:

1. If an uncompressed motion imagery signal contains a Precision Time Stamp in the VANC (Vertical Ancillary Data Space) (MISB STD 0605 [5]), the time stamp **shall** be extracted from the VANC and used to generate the Precision and Commercial Time Stamps for insertion into the motion imagery elementary stream (ES).
2. Otherwise, if an uncompressed motion imagery signal contains a Commercial Time Stamp in the VITC (Vertical Interval Time Code) or the VANC (Vertical Ancillary Data Space) (MISB STD 0605 [5]), the time stamp **shall** be extracted from the VITC or VANC and used to generate the Precision and Commercial Time Stamps for insertion into the motion imagery elementary stream (ES).
3. Otherwise, if the uncompressed motion imagery does not contain either time stamp, and a UTC referenced time source coordinated to the imagery is available, then that time source **shall** be used to generate the Precision and Commercial Time Stamps for insertion into the motion imagery elementary stream (ES).
4. Otherwise, if a UTC referenced time source is not available then a derived time stamp may be used to generate the Precision and Commercial Time Stamps for insertion into the

motion imagery stream and the value of Byte 17, bit 7 in Table 5-2 **shall** be set to one (indicating UTC lock unknown).

It is also **recommended** that both the Commercial Time Stamp and the Precision Time Stamp be inserted as early in the stream as possible.

5.2 MPEG-2 Compressed Elementary Stream

5.2.1 Precision Time Stamp: User Data Field

MPEG-2 (ISO 13818-2 [6]) provides for user defined data to be inserted into a MPEG-2 compressed elementary stream in the *user data field* (*user_data_start_code* = 0xB2). The Precision Time Stamp **shall** be inserted into the MPEG-2 elementary stream *user data field* located between the picture header and the picture data, so that it relates to a specific frame of video. The Precision Time Stamp **shall** be inserted every video frame.

This Precision Time Stamp **shall** be derived from a UTC reference of sufficient accuracy, such as GPS, as described in §4 of [4] and **shall** be formatted as defined in this section.

The *user data field* **shall** consist of two sub-fields: 1) an identification string and 2) a microsecond time stamp. This composite field comprising 28 total bytes (Byte 1 is transmitted first and as a series of 8-bit integers) is as follows:

Identification String: 16 bytes (Bytes 1-16) that represents the ASCII string "MISPmicrosectime" **shall** be set to the values in Table 5-1:

Byte 1	Byte 2	Byte 3	Byte 4	Byte 5	Byte 6	Byte 7	Byte 8
M	I	S	P	m	i	c	r
0x4D	0x49	0x53	0x50	0x6D	0x69	0x63	0x72
Byte 9	Byte 10	Byte 11	Byte 12	Byte 13	Byte 14	Byte 15	Byte 16
o	s	e	c	t	i	m	e
0x6F	0x73	0x65	0x63	0x74	0x69	0x6D	0x65

Table 5-1: Byte Assignment for ASCII String MISPmicrosectime

Microsecond time stamp: 12 bytes that represent a time stamp status byte followed by a 64-bit (8 byte) POSIX "Microseconds since 1970" time with embedded Start Code Emulation Prevention bytes. In Table 5-2 the 8 byte time stamp's most significant byte is Byte 1 and its least significant byte is Byte 8.

Byte 17	Status
bit 7	0 = Locked to valid UTC reference (internal oscillator clock locked to UTC)
	1 = UTC Lock Unknown (internal oscillator clock not locked to UTC)

bit 6	0 = Normal (time incremented normally since last message)
	1 = Discontinuity (time has not incremented normally since last message)
bit 5	0 = Forward (Bit 6 = 1 indicates that the time jumped forward)
	1 = Reverse (Bit 6 = 1 indicates that the time jumped backwards)
bits 4-0	Reserved ('11111')
Bytes 18, 19	Byte 1 and 2 (Most significant bytes) of Time Stamp (microseconds)
Byte 20	*Start Code Emulation Prevention Byte (0xFF)*
Bytes 21,22	Byte 3 and 4 of Time Stamp (microseconds)
Byte 23	*Start Code Emulation Prevention Byte (0xFF)*
Byte 24, 25	Byte 5 and 6 of Time Stamp (microseconds)
Byte 26	*Start Code Emulation Prevention Byte (0xFF)*
Byte 27, 28	Byte 7 and 8 (Least significant bytes) of Time Stamp (microseconds)

Table 5-2: Byte Assignment for 64-bit UTC Time Stamp

5.2.2 Commercial Time Stamp: GOP Header Time Code

The MPEG-2 video layer includes the definition of a 25-bit field time code (*time_code*) within the "group of pictures" (GOP) header. This time code is of the form HH:MM:SS:FF in a format specified in [6, 8]. The Commercial Time Stamp **shall** be inserted in every GOP.

5.2.3 MPEG-2 Video Elementary Stream Editing Information

Additional information that may be carried in the user data area of a video elementary stream is described in SMPTE ST 328 [9]. One of the additional metadata elements is a 64-bit time code which complies with SMPTE ST 12-1 [8] and SMPTE ST 309 [10]. The time code represents the time that the frame was captured (HH:MM:SS:FF), and it contains a date as defined in [10]. This information is not required nor expected to be the stream, and it is **not recommended** for stream timing.

5.3 H.264 Compressed Elementary Stream

The H.264 standard ISO/IEC 14496-10 [11] similarly provides data fields to include a time stamp and time code in the H.264 compressed elementary stream. These time references, described below, are inserted into the Supplemental Enhancement Information (SEI) message portion of the H.264 elementary stream.

5.3.1 Precision Time Stamp: User Data Unregistered SEI Message Field

The H.264 standard allows user defined data that is associated with a particular compressed frame with the *user data unregistered SEI Message* field (*user_data_unregistered*).
A Precision Time Stamp **shall** be inserted into the H.264 elementary stream *user data unregistered SEI Message* field, so that it relates to a specific frame of video.

The Precision Time Stamp **shall** be inserted every video frame.

This Precision Time Stamp **shall** be derived from a UTC reference of sufficient accuracy, such as GPS, as described in §4 of [4] and **shall** be formatted as defined in this section.

The *user data unregistered SEI message* **shall** consist of two sub-fields: 1) an identification string and 2) a microsecond time stamp. This composite field comprising 28 total bytes (Byte 1 is transmitted first) is as follows:

Identification String: The *uuid_iso_iec_11578* is a 16-byte field (value set as a UUID) that represents the ASCII string "MISPmicrosectime" **shall** be set to the values shown in Table 5-1:

Microsecond time stamp: Th*e user_data_payload_byte* is a variable length field with the same 12 byte format that is defined in Table 5-2:

5.3.2 Commercial Time Code: pic_timing

The H.264 standard provides for an optional time code to be inserted into the "picture timing SEI message." The picture timing SEI message (*pic_timing*) specifies this time code format as HH:MM:SS:FF. The time code reflects the time of frame capture, contains flags to specify whether the video is drop-frame, and whether there is a discontinuity in the video time-line. The Commercial Time Stamp **shall** be inserted in every video frame.

6 Time Stamping Metadata

Careful diligence in time stamping both the motion imagery and the metadata streams provides great value in future use of the asset. Moreover, should the two essence component streams become separated they can always be retimed together for display and processing.

MISB Standard 0603 [4] and the individual metadata standards **shall** be used for time stamping metadata.

7 Multiplexing of Motion Imagery and Metadata in MPEG-2 Transport Stream

Adding an absolute time stamp to motion imagery (§5.2 and §5.3 above) and metadata provides the means for correlating a motion imagery frame with metadata that is about the content within that motion imagery frame. The relationship between the motion imagery and the metadata is determined by these absolute time stamps, and is rather independent of the method selected for packaging the essence streams within an MPEG-2 transport stream. Two methods for packaging metadata along with motion imagery within an MPEG-2 Transport Stream are defined by ISO/IEC 13818-1 [1] and SMPTE RP 217 [13]; these are denoted here as the *synchronous metadata multiplex method* and the *asynchronous metadata multiplex method*, respectively.

7.1 Methods for Metadata Carriage

Synchronous metadata multiplex method

The ***synchronous* metadata multiplex method** ISO/IEC 13818-1 [1] provides for a PTS (Presentation Time Stamp) for metadata, similar to the PTS for the motion imagery. This affords precise alignment of the motion imagery with the metadata at the presentation stage—such as viewing on a display.

The PTS is a relative timing reference used to provide stream-to-stream synchronization within an MPEG-2 transport stream. Its use is principally to guarantee that each stream with the same PTS is presented (displayed) at the same time. It is not, however, intended to convey absolute time references; that is the function of the motion imagery and metadata absolute (UTC) time stamps.

ISO/IEC 13818-1 defines a metadata header structure which contains the PTS and other information described in §7.2.1 below. One important header field is a service ID that affords the labeling of a particular piece of metadata. As additional metadata is added downstream, it can be assigned its own service ID, thereby adding flexibility for segmenting metadata based on what elements are added, what corresponding motion imagery elementary stream the metadata applies, priority of the metadata in decoding, etc. However, until the usage of the service ID is defined the only allowed value **shall** be 0x00.

Asynchronous metadata multiplex method

The ***asynchronous* metadata multiplex method** SMPTE RP 217 [13] does not support a PTS for metadata, so that metadata cannot be guaranteed to be displayed co-incident to a particular frame of the motion imagery. Another difference is that unlike the synchronous method of carriage, where metadata can be assigned a unique ID, only one asynchronous stream is permitted per program in the transport stream. In more complex signal cases, for example, where multiple motion imagery streams are carried within a single program assigning different metadata to a specific motion imagery stream is not possible.

7.2 Requirements for Metadata Carriage

7.2.1 Synchronous Metadata Multiplex Method

Several methods to carry metadata within an MPEG-2 transport stream are detailed in ISO/IEC 13818-1 [1]. This MISB Standard requires the use of the method outlined in §2.12.4 “Use of PES packets to transport metadata” for transporting metadata that is synchronized with the essence stream, herein referred to as the **synchronous metadata multiplex method**. This method provides a means to synchronize essence streams using the Presentation Time Stamp (PTS) located in the Packetized Elementary Stream (PES) header. The PTS is coded into the MPEG-2 Systems PES layer, and is relevant for both MPEG-2 and H.264.

When metadata is sampled at the same time as a motion imagery frame the metadata and motion imagery frame will have the same PTS. If the metadata is not sampled at the same time as the motion imagery frame, it will have a different PTS, but will exist on the same timeline as the motion imagery frame—perhaps offset in time or at some non-periodic rate.

Figure 7-1 shows the structure of a PES packet in the metadata bit stream. In the most common implementations, the packet payload will consist of a single metadata Access Unit or *Metadata_AU_cell*, which includes a five-byte header followed by the KLV metadata.

A metadata service is defined in [1] as “a coherent set of metadata of the same format delivered to a receiver for a specific purpose… each metadata service is assumed to represent a concatenation (or collection) of metadata Access Units.” Each metadata service is represented by a collection of metadata Access Units which are transported in PES packets. A metadata Access Unit Cell contains a 5-byte header (structured as in Table 7-1).

When transporting metadata using this service, a unique *metadata_service_id* is assigned to each service.

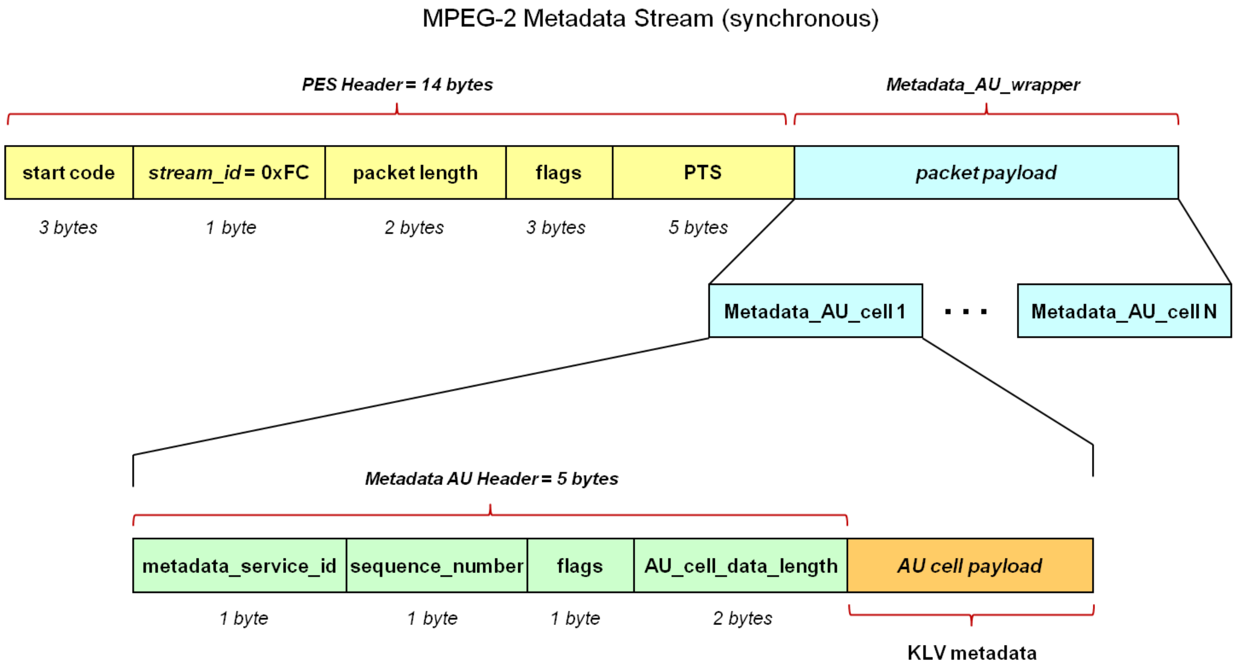

Figure 7-1: Synchronous Metadata Multiplex Method

Requirements for synchronous metadata multiplexing are:

<u>Synchronous Metadata PES Stream</u>

1. Insertion of synchronous metadata **shall** use ISO 13818-1 §2.12.4 [1].
2. The *stream_id* **shall** be 0xFC, indicating “metadata stream”.
3. Each PES packet **shall** have a PTS to be used to synchronize the metadata with the motion imagery. *PTS_DTS_flags* **shall** be set to ‘10’.
4. In each PES packet that carries metadata, the first PES packet data byte **shall** be the first byte of a Metadata Access Unit Cell.
5. The PTS in the PES header **shall** apply to each Access Unit contained in the PES packet.

6. The PTS **shall** signal the time that the metadata Access Unit becomes relevant. It is assumed that the metadata is decoded instantaneously (i.e., no DTS **shall** be coded). If a motion imagery frame and a metadata Access Unit have the same PTS, then they were sampled at the same time.
7. Each metadata Access Unit **may** be carried in one or more Access Unit Cells.
8. The delay of any data through the System Target Decoder buffers **shall** be less than or equal to one second.
9. When inserting synchronous metadata into a transport stream which already carries synchronous metadata new metadata **shall** be added to the existing synchronous metadata stream.
10. There **shall** be only one Packet ID (PID) for all collective services per program.
11. Newly added metadata services **shall** use only *metadata_service_id* = 0x00.

<u>Program Map Table (PMT)</u>

1. The stream_type **shall** be 0x15, indicating "Metadata carried in PES packets".
2. The Metadata Stream **shall** be defined in the PMT as a separate stream within the same Program as the motion imagery elementary stream. Multi-program Transport Streams and methods for associating metadata in one program to motion imagery in another are allowed [1]. Multi-program Transport Streams are <u>not</u> covered within the scope of this STANDARD.
3. The PMT **shall** contain a *metadata_descriptor* for each metadata service within the metadata stream. The *metadata_descriptor* **shall** be within the descriptor loop for the metadata stream. The *metadata_descriptor* contains the *metadata_service_id* for the service it describes.
4. The PMT **shall** contain a single *metadata_std_descriptor* for the metadata stream.
5. The PMT **may** contain other descriptors such as the *content_labeling_descriptor* and the *metadata_pointer_descriptor*.

Table 7-1 shows sample values for the *metadata_descriptor*, *metadata_std_descriptor* and *metadata_AU_cell* header fields when using constructing a synchronous metadata stream.

Table 7-1 – Sample Metadata Descriptors for Synchronous Metadata

	Value	No. of bits
metadata_descriptor		
descriptor_tag	0x26 (38)	8
descriptor_length	0x09 (9)	8
metadata_application_format	0x0100 - 0x0103 (see Table 7-3)	16
metadata_format	0xFF	8
metadata_format_identifier	0x4B4C5641 = "KLVA"	32
metadata_service_id	0x00	8
decoder_config_flags	'000'	3

DSM-CC_flag	'0'	1
reserved	'1111'	4
metadata_std_descriptor		
descriptor_tag	0x27 (39)	8
descriptor_length	0x09 (9)	8
reserved	'11'	2
metadata_input_leak_rate	(determined by encoder)	
reserved	'11'	2
metadata_buffer_size	(determined by encoder)	
reserved	'11'	2
*metadata_output_leak_rate**	(unspecified; recommend setting to 0)	
Metadata_AU_cell (5-byte header)		
metadata_service_id	0x00	8
sequence_number	(supplied by encoder; increments each cell)	8
cell_fragmentation_indication	'11', '10', '01' or '00'	2
decoder_config_flag	'0'	1
random_access_indicator	'0' or '1'	1
reserved	'1111'	4
AU_cell_data_length	(supplied by encoder)	16

*NOTE: the *metadata_output_leak_rate* is unspecified for synchronous metadata. The recommended value is 0.

7.2.2 Asynchronous Metadata Multiplex Method

The transport of KLV metadata over an MPEG-2 transport stream in an asynchronous manner **shall** be confined to the method defined in SMPTE RP 217 §4.1.1 or §4.1.2 [13]. As shown in Figure 7-2, the metadata PES packets do not employ a Presentation Time Stamps (PTS) or Metadata Access Unit construct. The time relationship between the metadata and a motion imagery frame is established by proximity.

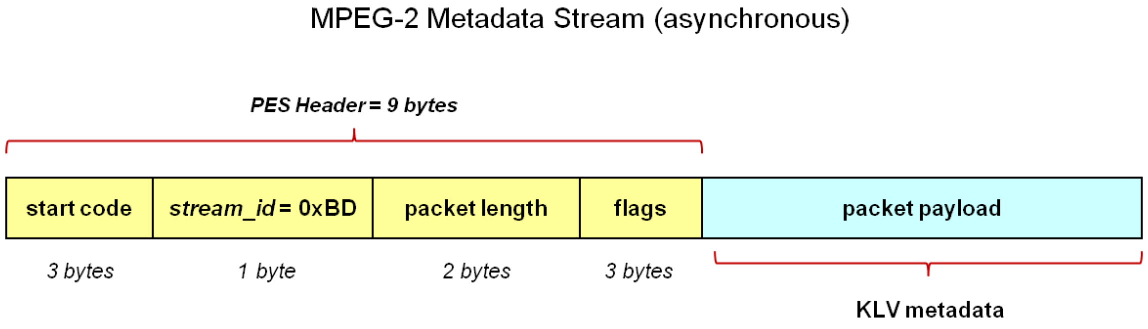

Figure 7-2: Asynchronous Metadata Multiplex Method

Requirements for asynchronous metadata multiplexing are:

Asynchronous Metadata PES Stream

1. The *stream_id* **shall** be 0xBD, indicating “private_stream_1.”
2. The *data_alignment_indicator* **shall** be set to one when the PES packet contains the beginning of a KLV item, and **shall** be set to zero otherwise.
3. The *PTS_DTS_flags* **shall** be set to 00 (no PTS or DTS present in PES packet header).
4. There **shall** be only one Packet ID (PID) for private_stream_1 per program.
5. When inserting asynchronous metadata into a transport stream which already carries asynchronous metadata of the same *format_identifier*, it **shall** be inserted into the existing metadata elementary stream.

Program Map Table (PMT)

1. The *stream_type* **shall** be 0x06, indicating “PES packets containing private data.”
2. The metadata stream **shall** be defined in the PMT as a separate Stream within the same Program as the motion imagery elementary stream. Multi-program Transport Streams and methods for associating metadata in one program to video in another are allowed [1]. Multi-program Transport Streams are not covered within the scope of this standard.
3. For legacy compliance with [13], the program element loop in the PMT **shall** contain a *registration_descriptor* as defined in [1].
4. The PMT **shall** contain a *metadata_descriptor* for each metadata service within the metadata stream. The *metadata_descriptor* **shall** be within the descriptor loop for the metadata stream. The *metadata_descriptor* contains the *metadata_service_id* for the service it describes. NOTE: Earlier versions of [1] describe the use of the *registration_descriptor* to “uniquely and unambiguously identify formats of private data.” The *metadata_descriptor*, however, provides more functionality, and is therefore specified.
5. The *metadata_service_id* **shall** be unique from other services within a program.
6. The PMT **shall** contain a single *metadata_std_descriptor* for the metadata stream.
7. The PMT **may** contain other descriptors such as the *content_labeling_descriptor* and the *metadata_pointer_descriptor*.

Table 7-2 shows sample values for the *registration_descriptor*, *metadata_descriptor* and *metadata_std_descriptor* when constructing an asynchronous metadata stream.

Table 7-2 – Sample Descriptors for an Asynchronous Metadata Stream

	Value	No. of bits
registration_descriptor		
descriptor_tag	0x05 (5)	8
descriptor_length	0x04 (4)	8
format_identifier	0x4B4C5641 = “KLVA”	32
metadata_descriptor		

descriptor_tag	0x26 (38)	8
descriptor_length	0x09 (9)	8
metadata_application_format	0x0100 - 0x0103 (see Table 7-3)	16
metadata_format	0xFF	8
metadata_format_identifier	0x4B4C5641 = “KLVA”	32
metadata_service_id	0x00	8
decoder_config_flags	‘000’	3
DSM-CC_flag	‘0’	1
reserved	‘1111’	4
metadata_std_descriptor		
descriptor_tag	0x27 (39)	8
descriptor_length	0x09 (9)	8
reserved	‘11’	2
metadata_input_leak_rate	(determined by encoder)	
reserved	‘11’	2
metadata_buffer_size	(determined by encoder)	
reserved	‘11’	2
*metadata_output_leak_rate**	(determined by encoder)	

*NOTE: the *metadata_output_leak_rate* must be specified for asynchronous metadata.

Table 7-3 – KLV Metadata Type

metadata_application_format (type of KLV metadata)	
0x0100 (default)	General
0x0101	Geographic Metadata
0x0102	Annotation Metadata
0x0103	Still Image on Demand

Appendix A - Calculation of Presentation Time Stamp (PTS) - *Informative*

The following information is only valid for systems where the motion imagery source clock is genlocked to the time stamp clock source, such as that provided by GPS. If this is not the case, the PTS and UTC will drift apart over time**.**

Metadata "sampling accuracy" should be a function of the time stamp applied to the metadata stream itself and not determined by the PTS. However, forming an accurate PTS will aide in presenting/displaying the motion imagery with the metadata, which is extremely critical for viewing.

It is **recommended** that inter-stream essence PTS values reflect the actual difference in the UTC time stamps of the respective essence streams. In general, a system that produces a UTC time stamp for metadata and a UTC time stamp for motion imagery that are equivalent should produce a corresponding PTS for the metadata equivalent to the PTS for the motion imagery.

Figure A-1 depicts two time bases: UTC and an MPEG-2 TS Encoder System Time Clock (STC) that forms the basis of inter-stream timing. **Frame 1** and **Frame 2** of motion imagery occur at a periodic temporal rate (e.g. 30, 60 Hz) and are time stamped with UTC times $T_{V1\text{-}UTC}$, $T_{V2\text{-}UTC}$, etc., while **Metadata** is similarly time stamped with $T_{M1\text{-}UTC}$ on the UTC time line. The MPEG-2 transport stream clock, a relative timing system, provides a PTS value $T_{V1\text{-}PTS}$, $T_{V2\text{-}PTS}$ for each of the motion imagery frames, and $T_{M1\text{-}PTS}$ for the metadata on the STC time line.

The difference in time between **Frame 1** and **Metadata** is the same for both time scales. The PTS for the **Metadata** can then be computed from the PTS of **Frame 1** and either one of the timing difference signals—ΔPTS or ΔUTC as shown in the figure. Formulas that convert absolute time and time code to PTS values can be found in SMPTE EG 40 [14], which should provide an implementer the ability to form PTS values that are very consistent with the UTC time stamps within motion imagery and metadata streams.

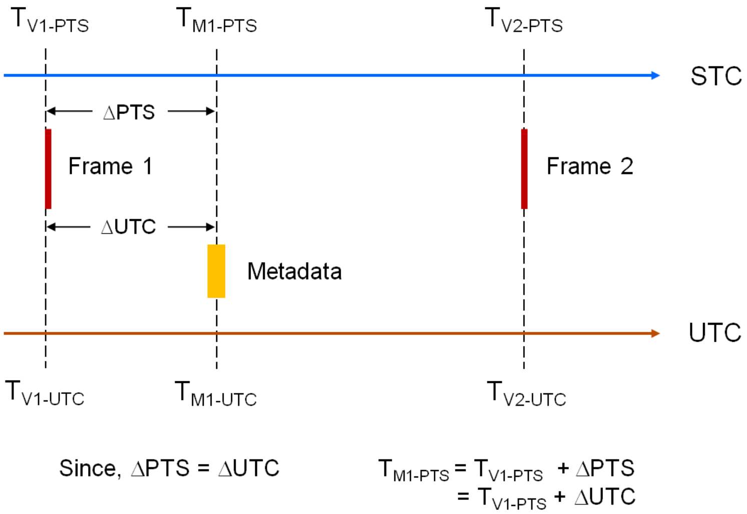

Figure A-1: TS System Clock and UTC Relationship